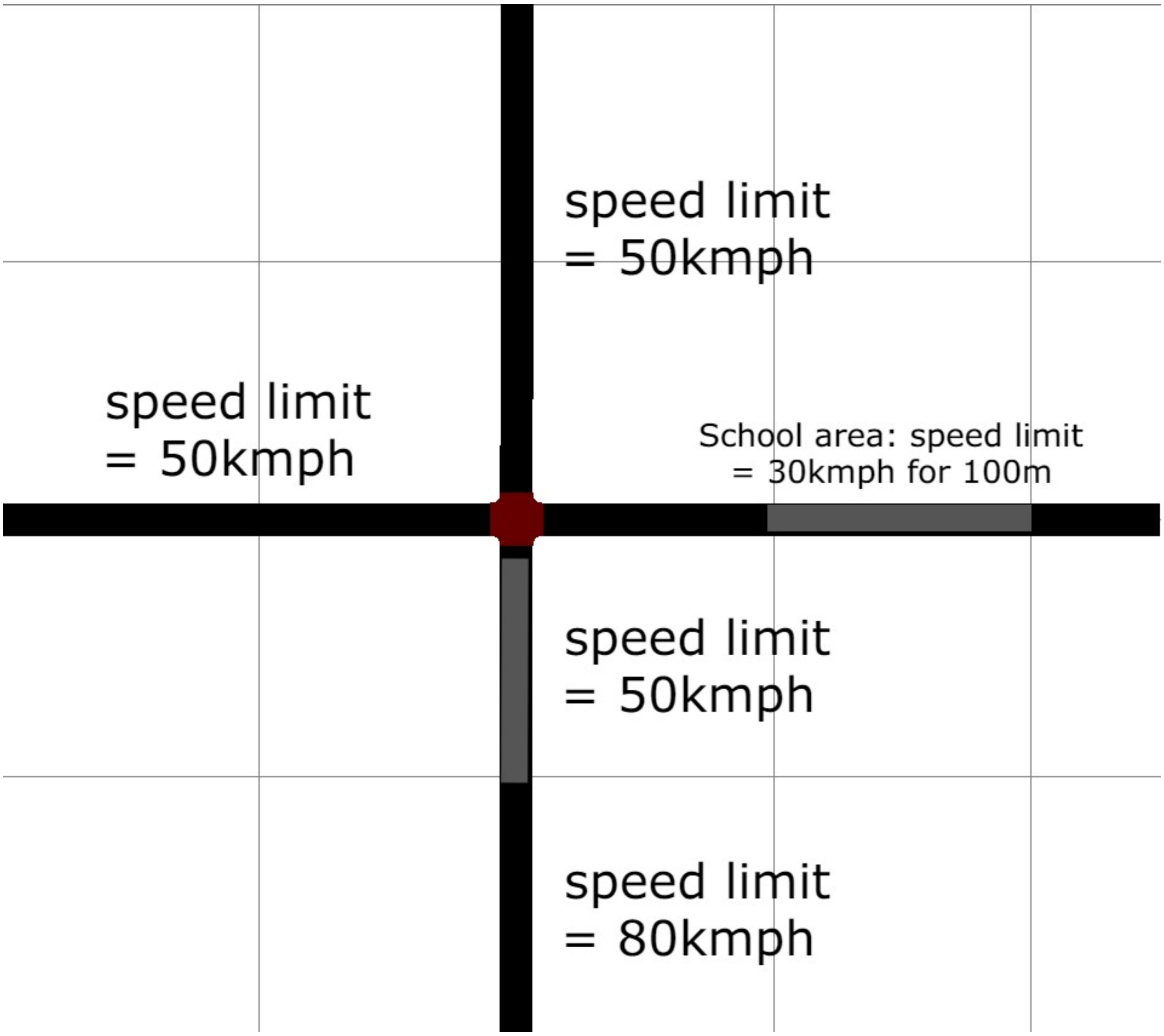
Weekly Updates (7/12/23) – Vinayak Gajendra Panchal

Ego vehicle collision prediction – LSTM (nearest_vehicle_distance approach)

Simulation setup:

Step-length	0.10 sec (10 cycles/sec)
Collision.action	remove
Collision.stoptime	2 sec
Collision.check-junctions	True
'- - log' (simulation log files)	simulation_main.log
Number of Vehicles	11 (1 ego-vehicle, 10 non-ego vehicle)

2-Lane 4-Way Junction Details:



Dataset Creation: I have integrated lanes with speed limits, including a school zone with a maximum speed of 30 km/h (marked in grey), to better replicate traffic scenarios in the LaneLet2 framework. In this simulation, 11 vehicles were modeled, with data collected every 0.1 seconds across three different sets (training, validation, and testing). I prepared a total of 10 datasets, 5 with collision attributes (collision happened) and 5 without collisions. All these datasets were prepared with ego-vehicle as the main character for collision predictions in the next k steps.

Dataset features (for all train-validation-test datasets):

1. **Time:** Time in seconds when the ego vehicle was present in the simulation.
2. **ego-speed, ego-acceleration:** The acceleration and speed value of ego-vehicle at all timestamps
3. **ego_x, ego_y:** spatial data of ego-vehicle at all timestamps.
4. **nearest_vehicle_id:** nearest_vehicle id based on the nearest Euclidean distance of non-ego vehicle with ego-vehicle.
5. **nearest_vehicle_speed:** Calculated speeds of nearest non-ego vehicles at each timestamp.
6. **nearest_vehicle_acceleration:** Acceleration of the nearest non-ego vehicles at each timestamp.
7. **nearest_vehicle_relative_displacement:** Calculated relative displacement (Euclidian distance) to nearest non-ego vehicles at each timestamp with the ego-vehicle.

Trend in Datasets (nearest_vehicle_relative_displacement):

The "Simulation with Collision" plot tends to show a decrease to lower values followed by stabilization, which reflects a decrease in variables like speed or distance post-collision. The "Simulation without Collision" plot shows more fluctuation, indicating ongoing variability in relative displacement values, because the variables continue to change freely without the interruption of a collision.

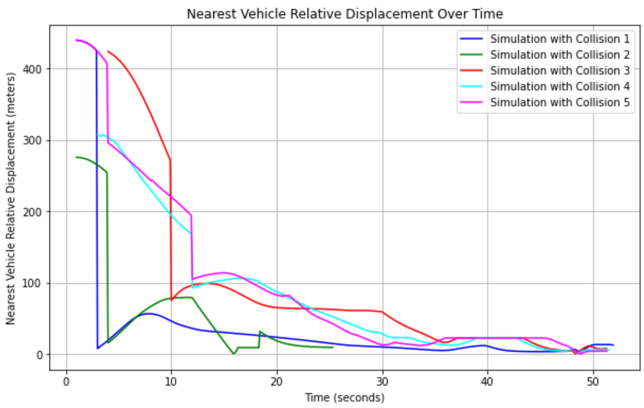


Figure 1: Nearest Displacement – simulation with collision

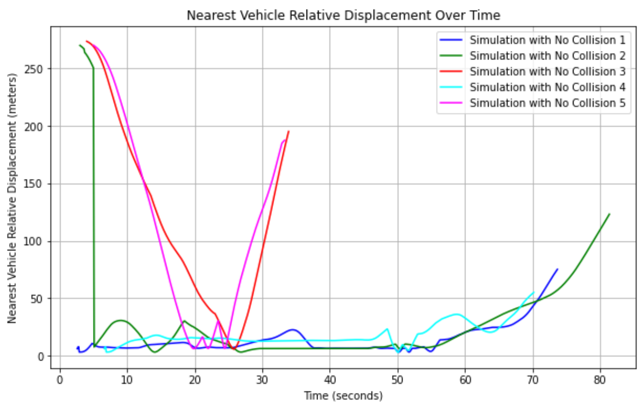


Figure 2: Nearest Displacement – simulation with no collision

LSTM Model: An LSTM (Long Short-Term Memory) architecture was employed to forecast the upcoming 5 timestamp rows, including variables like ego_speed, ego_acceleration, ego_x, ego_y, nearest_vehicle_acceleration, nearest_vehicle_speed, and nearest_vehicle_relative_ displacement, by analyzing the preceding 15 rows. The LSTM architecture used for training these datasets involves 2 LSTM layers one with 100 units and the other with 50, 1 dense layer with 175 neurons, and a final layer with 35 neurons (5 X 7 features).

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #   
=================================================================
 lstm (LSTM)                 (None, 15, 100)           43200     
                                                                 
 lstm_1 (LSTM)               (None, 50)                30200     
                                                                 
 dense (Dense)               (None, 175)               8925      
                                                                 
 dropout (Dropout)           (None, 175)               0         
                                                                 
 dense_1 (Dense)             (None, 35)                6160      
                                                                 
=================================================================
Total params: 88485 (345.64 KB)
Trainable params: 88485 (345.64 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Actual 5 timestamp rows to be predicted:

	ego_speed	ego_acceleration	ego_x	ego_y	nearest_vehicle_speed	nearest_vehicle_acceleration	nearest_vehicle_relative_displacement
455	7.27	2.54	0.89	-2.24	5.09	-4.50	2.023858
456	7.37	1.05	0.51	-2.88	4.65	-4.48	1.668323
457	7.51	1.35	0.13	-3.52	4.26	-3.88	1.938891
458	7.67	1.65	-0.27	-4.18	3.93	-3.28	2.631881
459	6.67	-10.00	-0.61	-4.75	3.66	-2.68	3.381627

Study 1: Focusing on a single dataset from the simulation with the occurrence of a collision.

- Both training and test losses decrease sharply at the beginning, indicating quick learning in the initial epochs.
- The actual and predicted timesteps show a large discrepancy in the ego_x, ego_y, acceleration values, and nearest_vehicle_relative_displacement which is reflected in the higher MSE for these variables.
- The high MSE for ego_x suggests a significantly large prediction error in the vehicle position.
- The prediction was not accurate, no rows matched the actual.

Columns	Ego_speed	Ego_accel	Ego_x	Ego_y	Near_veh_speed	Near_veh_accel	Relative displacement
MSE	4.1477	22.8985	1530.9983	40.9639	5.9270	17.4495	43.1791

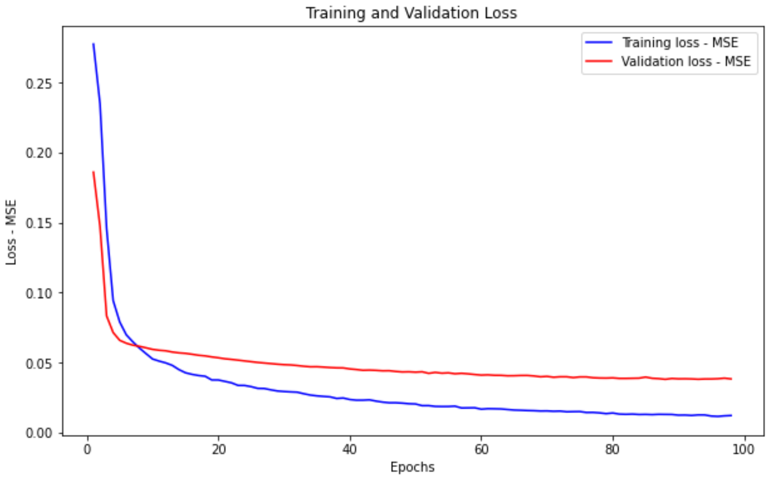


Figure 3: Mean Square Error loss of both training and validation for study 1

Predicted 5 timestamp rows:

	ego_speed	ego_acceleration	ego_x	ego_y	nearest_vehicle_speed	nearest_vehicle_acceleration	nearest_vehicle_relative_displacement
0	5.530689	0.232990	29.735617	3.260863	6.966780	0.427177	-1.667304
1	5.679511	0.085402	40.785477	2.310704	6.867776	0.539948	8.562242
2	4.875995	-0.034054	43.300751	3.080216	6.802257	0.233234	13.686946
3	5.323971	-0.256757	45.390308	2.509929	6.707571	0.063324	6.587652
4	5.116354	0.133540	34.776871	2.950565	6.308828	0.430363	4.420842

Study 2: Focusing on all 5 datasets from the simulation with the occurrence of a collision.

- The training and validation loss curves show more variability and fluctuations than in Study 1.
- The lower MSE for nearest_vehicle_speed suggests that the model predicts the speed of the nearest vehicle more accurately than in Study 1.
- Study 2 predicts ego_x better than study 1. The nearest relative displacement remains the same for both studies 1 and 2.

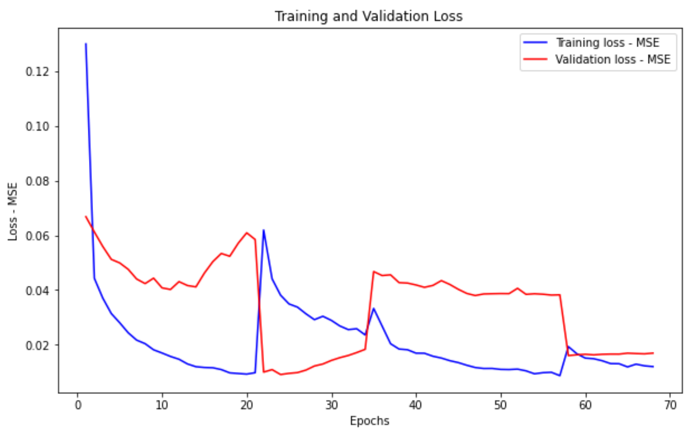


Figure 4: Mean Square Error loss of both training and validation for study 2

Columns	Ego_speed	Ego_accel	Ego_x	Ego_y	Near_veh_speed	Near_veh_accel	Relative displacement
MSE	15.8316	23.2951	20.3010	16.5789	2.8831	17.8773	43.6655

Predicted 5 timestamp rows:

	ego_speed	ego_acceleration	ego_x	ego_y	nearest_vehicle_speed	nearest_vehicle_acceleration	nearest_vehicle_relative_displacement
0	3.216607	-1.767124	-5.701119	0.510598	6.093621	0.451250	6.336795
1	3.426393	-1.992658	-2.453146	0.459176	6.007215	0.481302	12.044146
2	3.317338	-1.890387	-4.849128	0.432850	6.005793	0.401064	10.212312
3	3.475820	-2.089043	-4.808394	0.535305	5.904314	0.332700	7.373904
4	3.204973	-1.988391	-2.583175	0.380252	5.809485	0.299588	4.445731

Study 3: Considering all 5 datasets from the simulation with the occurrence of a collision.

- The Below plot shows an extreme spike in the training loss, indicating a possible anomaly or drastic change in the dataset at that point from collision attributable to no collision.
- The 5 predicted timesteps show significant variance from the actual timesteps, with the relative displacement and ego_accelaration having higher MSE values compared to studies 1 and 2, indicating poor prediction performance. Predicted great ego_speed with low 0.4 MSE.

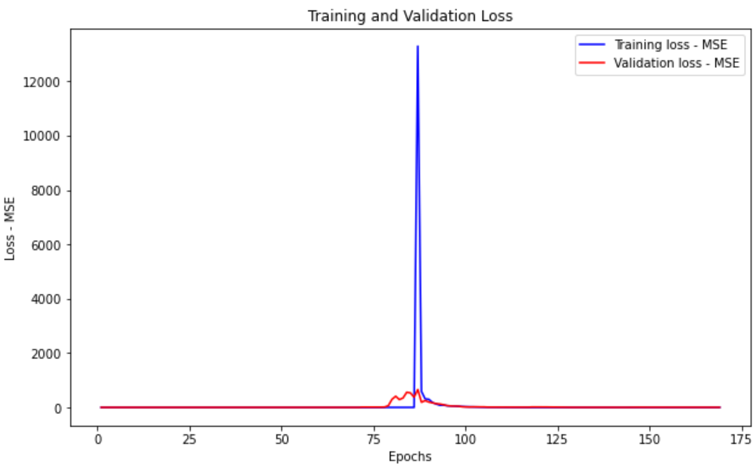


Figure 4: Mean Square Error loss of both training and validation for study 3

Predicted 5 timestamp rows:

	ego_speed	ego_acceleration	ego_x	ego_y	nearest_vehicle_speed	nearest_vehicle_acceleration	nearest_vehicle_relative_displacement
0	7.011085	-2.701354	3.553900	-5.337836	8.323498	-0.056874	7.447763
1	6.586177	-2.111078	19.834400	-6.061495	7.948454	-0.224143	23.322729
2	7.110813	-2.845597	4.552804	-6.721208	8.732163	-0.294258	33.649212
3	6.909030	-1.927503	-0.765039	-6.876007	8.786139	-0.641894	13.684005
4	7.568554	-2.163022	11.788939	-7.028240	7.982291	-0.590981	20.771879

Columns	Ego_speed	Ego_accel	Ego_x	Ego_y	Near_veh_speed	Near_veh_accel	Relative displacement
MSE	0.4454	25.8568	110.8138	8.4850	16.7199	12.4070	385.6894

Comparing Studies: Study 1 shows the most consistent and stable training and validation losses. Study 2 shows more fluctuation in losses than study 1 but less compared to study 3, which may be due to model exposure to multiple datasets. The prediction accuracy, based on MSE values, is best in Study 1, while Study 3 shows the least accurate predictions. Studies 2 and 3 overall show a decrease in prediction performance, with increased MSE values across most variables.

In conclusion, when comparing the three studies, it appears that using a single, consistent dataset for training leads to more stable and accurate predictions for the nearest_vehicle_distance approach. The presence of anomalies or high variability across multiple datasets can introduce learning challenges and result in higher prediction errors. Also considering nearest displacement with ego vehicles is not a good way to predict collisions reason as vehicles can be nearby even when passing by the lanes, standing in traffic, etc. Also based on the inaccuracies in predicting the nearest distance we can not use this approach to predict collisions.